# Mixture Models in SEM

Petra Stein
Universität Duisburg-Essen
petra_stein@uni-duisburg-essen.de

Structural equation modelling, or generally speaking mean and covariance structure models, are a well-tried statistical approach to analyzing simultaneously relations between unobserved variables (structural equation model), and relations between unobserved variables and observed variables (measurement model). Normally, the underlying mean and covariance structure is assumed to stem from one population. This assumption is often unrealistic. If heterogeneity exists and the variable which generates the heterogeneity is known, the researcher may use the multiple group options in programs like LISREL or EQS to estimate the parameter of the structural and the measurement model in different groups, and test the model.

A central topic of empirical research is the problem of unobserved heterogeneity. To solve this problem at least partially, a statistical model, the so-called mixture model, is developed. "Finite mixtures" refers to the assumption that a sample of observations arises from a mixture of normal, exponential, or Bernoulli distribution. Finite mixtures of multivariate normal variables may be used whenever the researcher suspects that a population is not homogeneous but consists of heterogeneous subpopulations that have different means and associations between the variables of each subpopulation. In addition to describing each component, the researcher may want to specify mean and covariance structure models like the LISREL models for each component, and may also want to estimate the size of each component. Finite mixtures may further be used as an alternative to cluster analytic techniques in order to identify homogeneous subsets in samples from a heterogeneous population, in the sense that elements within the same subset follow a multivariate normal distribution with the same expected value and variance matrix.

Further developments of the mixture model are conditional finite mixtures. In conditional mixtures of multivariate normal distributions, the expected value of each component may depend linearly on regressor variables that need not be multivariate normal. Consequently, the strong assumption of unconditional normality of the mixture components is replaced by the weaker assumption of conditional normality. Further developments of unconditional finite mixtures are the integration of full path models or confirmatory factor models in the mixtures, a development of conditional finite mixtures is a full general mean and covariance model. Another development involves longitudinal data in the model, like for example the growth mixture model.

In this session we want to consider some of the new and different developments and applications of mixture models in structural equation modelling.